A Semantic Network Dictionary for Dysphasia Therapy

Bácsi János

bacsi@jgytf.u-szeged.hu

Keywords: delayed speech development, frequency dictionary, network theories, semantic structure, machine-aided speech therapy

The purpose of this project is to compile a network dictionary running on computers to help the conscious planning of the speech therapy of 4–8 year-old-children suffering from expressive language disorder. The children will be able to take a virtual tour through the connected words of the network dictionary only by saying words into a microphone, by which the program will help in the development and maintenance of semantic structures.

Our basic psychological assumption, reinforced by our test results, is that a call word that enters the STM (short-term memory) will retrieve about 2–5 concepts from the LTM (long-term memory) and its association with them produces the semantic structure of the call word.

One of our linguistic achievements is that we have compiled a frequency dictionary based on thirteen primers. It contains 27,293 grapheme sequences of which 12,226 present content words. The goal of the dictionary is to find out which are the words that a six-year-old child are most likely to come across during learning the written language. The empirical database of the network dictionary will be drawn by asking 5000 children aged 4–7 to tell us what comes into their minds when they hear the 200 most frequently used nouns of the frequency dictionary. We have done half of that work so far. The five more frequent associations prompted by the word *anya* 'mother,' which is among the words that have been completely processed, are *szeret* 'she loves me' 1588, *szülő* 'parent' 1501, *szeretet* 'love' (noun) 1127, *szeretem* 'I love her' 875. The number of the distinct associations prompted by 'mother' is 741. We also have a large database of segmented and annotated recordings of children's voice, which contains 250,000 items.

The computational task to produce a program for the network dictionary that uses speech recognition has already been accomplished. The other task is to create a network dictionary that demonstrates all of the possible associations based on the empirical material that has accumulated 400,000 words so far.

The expected result makes it predictable what are the most like associations evoked by certain concepts among children aged 4–7, which makes the therapy of delayed speech development plannable using the words that the children have already learnt.